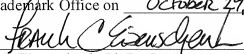


I hereby certify that this correspondence is being electronically filed in the United States Patent and Trademark Office on October 29, 2007.



Frank C. Eisenschenk, Ph.D., Patent Attorney

REQUEST FOR CERTIFICATE OF  
CORRECTION UNDER 37 CFR 1.322  
Docket No. GEN-T111XC3D2  
Patent No. 7,189,833

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants : Daniel Cohen, Marta Blumenfeld, Ilya Chumakov, Lydie Bougueleret  
Issued : March 13, 2007  
Patent No. : 7,189,833  
For : Prostate Cancer Gene

Mail Stop Certificate of Corrections Branch  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

REQUEST FOR CERTIFICATE OF CORRECTION  
UNDER 37 CFR 1.322 (OFFICE MISTAKE)

Sir:

A Certificate of Correction for the above-identified patent has been prepared and is attached hereto.

In the left-hand column below is the column and line number where errors occurred in the patent. In the right-hand column is the page and line number in the application where the correct information appears.

**Patent Reads:**

Column 19, line 12:

“The term”

Column 19, line 53:

“and 35”

**Application Reads:**

Page 24, line 20:

--The term--

Page 25, line 16:

--and 315--

Column 22, line 22:

“Washburn et al.”

Column 22, lines 24-25:

“(Washburn”

Column 29, line 28:

“5 MM MgC1<sub>2</sub>”

Column 30, line 25:

“72° C. final elongation”

Column 31, line 45:

“The re was incubated”

Column 47, line 14:

“Tumor Suppressor”

Column 51, line 10:

“44-187”

Column 51, line 12:

“442-304”

Column 51, line 13:

“443-328”

Column 51, line 28:

“SEQ 1) NOs 464 and”

Column 51, line 54:

“99-221442 of”

Page 29, line 3:

--Washburn et al.--

Page 29, line 4:

--(Washburn--

Page 38, line 9:

--5 mM MgC1<sub>2</sub>--

Page 39, lines 16-17:

--72° C. For final elongation--

Page 41, lines 4-5:

--The reaction product was incubated--

Page 61, line 4:

-- Tumor Suppressor--

Page 66, line 14:

-- 4-4-187--

Page 66, line 15:

-- 4-42-304--

Page 66, line 16:

-- 4-43-328--

Page 66, line 27:

--SEQ ID NOs 464 and--

Page 67, line 14:

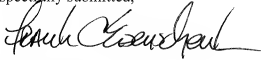
--99-221-442 of--

<u>Column 82, line 37:</u>	<u>Page 106, line 15:</u>
“by loxp sites”	--by loxP sites--
<u>Column 84, line 6:</u>	<u>Page 108, line 16:</u>
“(Bemoist”	--(Bernoist--
<u>Column 85, line 4:</u>	<u>Page 109, line 21:</u>
“(Stemberg N. L., 1994”	--(Sternberg N.L., 1994--
<u>Column 89, line 56:</u>	<u>Page 115, line 23:</u>
“the loxp sites”	--the loxP sites--
<u>Column 101, lines 59-60:</u>	<u>Page 130, line 13:</u>
“with 0-phenylenediamine”	--with o-phenylenediamine--
<u>Column 123, line 9:</u>	<u>Page 157, line 5:</u>
“nucteic acid”	--nucleic acid--.

A true and correct copy of pages 24, 25, 29, 38, 39, 41, 60, 61, 66, 67, 106, 108, 109, 115, 130, and 157 of the specification as filed which support Applicants' assertion of the errors on the part of the Patent Office accompanies this Certificate of Correction.

Approval of the Certificate of Correction is respectfully requested.

Respectfully submitted,



Frank C. Eisenschenk, Ph.D.

Patent Attorney

Registration No. 45,332

Phone No.: 352-375-8100

Fax No.: 352-372-5800

Address: P.O. Box 142950

Gainesville, FL 32614-2950

FCE/dc/sl

Attachments: Copy of pages 24, 25, 29, 38, 39, 41, 60, 61, 66, 67, 106, 108, 109, 115, 130, and 157 of the specification

5-56-225 A or G in position 50054; 5-56-272 A or G in position 50101;  
 5-56-391 G or T in position 50220; 4-61-269 A or G in position 50440;  
 4-61-391 A or G in position 50562; 4-63-99 A or G in position 50653;  
 4-62-120 A or G in position 50660; 4-62-205 A or G in position 50745;  
 5 4-64-113 A or T in position 50885; 4-65-104 A or G in position 51249;  
 5-28-300 A or G in position 51333; 5-50-269 C or T in position 51435;  
 4-65-324 C or T in position 51468; 5-71-129 G or C in position 51515;  
 5-50-391 G or C in position 51557; 5-71-180 A or G in position 51566;  
 4-67-40 C or T in position 51632; 5-71-280 A or C in position 51666;  
 10 5-58-167 A or G in position 52016; 5-30-325 C or T in position 52096;  
 5-58-302 A or T in position 52151; 5-31-178 A or G in position 52282;  
 5-31-244 A or G in position 52348; 5-31-306 deletion of A in position 52410;  
 5-32-190 C or T in position 52524; 5-32-246 C or T in position 52580;  
 5-32-378 deletion of A in position 52712; 5-53-266 G or C in position 52772;  
 15 5-60-158 C or T in position 52860; 5-60-390 A or G in position 53092;  
 5-68-272 G or C in position 53272; 5-68-385 A or T in position 53389;  
 5-66-53 deletion of GA in position 53511; 5-66-142 G or C in position 53600;  
 5-66-207 A or G in position 53665; 5-37-294 A or G in position 53815;  
 5-62-163 insertion of A in position 54365; 5-62-340 A or T in position 54541; and the  
 20 compliments thereof. The term PG1-related biallelic marker also includes all of the following  
 biallelic markers listed by internal reference number, and two SEQ ID NOs each of which  
 contains a 47-mers with one of the two alternative bases at position 24:  
 4-14-107 of SEQ ID NOs 185 and 262; 4-14-317 of SEQ ID NOs 186 and 263; 4-14-35  
 of SEQ ID NOs 187 and 264; 4-20-149 of SEQ ID NOs 188 and 265;  
 25 4-20-77 of SEQ ID NOs 189 and 266; 4-22-174 of SEQ ID NOs 190 and 267;  
 4-22-176 of SEQ ID NOs 191 and 268; 4-26-60 of SEQ ID NOs 192 and 269;  
 4-26-72 of SEQ ID NOs 193 and 270; 4-3-130 of SEQ ID NOs 194 and 271;  
 4-38-63 of SEQ ID NOs 195 and 272;  
 4-38-83 of SEQ ID NOs 196 and 273; 4-4-152 of SEQ ID NOs 197 and 274;  
 30 4-4-187 of SEQ ID NOs 198 and 275; 4-4-288 of SEQ ID NOs 199 and 276;  
 4-42-304 of SEQ ID NOs 200 and 277; 4-42-401 of SEQ ID NOs 201 and 278;  
 4-43-328 of SEQ ID NOs 202 and 279; 4-43-70 of SEQ ID NOs 203 and 280;  
 4-50-209 of SEQ ID NOs 204 and 281; 4-50-293 of SEQ ID NOs 205 and 282;  
 4-50-323 of SEQ ID NOs 206 and 283; 4-50-329 of SEQ ID NOs 207 and 284;

4-50-330 of SEQ ID NOs 208 and 285; 4-52-163 of SEQ ID NOs 209 and 286;  
 4-52-88 of SEQ ID NOs 210 and 287; 4-53-258 of SEQ ID NOs 211 and 288;  
 4-54-283 of SEQ ID NOs 212 and 289; 4-54-388 of SEQ ID NOs 213 and 290;  
 4-55-70 of SEQ ID NOs 214 and 291; 4-55-95 of SEQ ID NOs 215 and 292;  
 5 4-56-159 of SEQ ID NOs 216 and 293; 4-56-213 of SEQ ID NOs 217 and 294;  
 4-58-289 of SEQ ID NOs 218 and 295; 4-58-318 of SEQ ID NOs 219 and 296;  
 4-60-266 of SEQ ID NOs 220 and 297; 4-60-293 of SEQ ID NOs 221 and 298;  
 4-84-241 of SEQ ID NOs 222 and 299; 4-84-262 of SEQ ID NOs 223 and 300;  
 4-86-206 of SEQ ID NOs 224 and 301; 4-86-309 of SEQ ID NOs 225 and 302;  
 10 4-88-349 of SEQ ID NOs 226 and 303; 4-89-87 of SEQ ID NOs 227 and 304;  
 99-123-184 of SEQ ID NOs 228 and 305; 99-128-202 of SEQ ID NOs 229 and 306;  
 99-128-275 of SEQ ID NOs 230 and 307; 99-128-313 of SEQ ID NOs 231 and 308; 99-  
 128-60 of SEQ ID NOs 232 and 309; 99-12907-295 of SEQ ID NOs 233 and 310; 99-130-58 of  
 SEQ ID NOs 234 and 311; 99-134-362 of SEQ ID NOs 235 and 312;  
 15 99-140-130 of SEQ ID NOs 236 and 313; 99-1462-238 of SEQ ID NOs 237 and 314; 99-147-  
 181 of SEQ ID NOs 238 and 315; 99-1474-156 of SEQ ID NOs 239 and 316; 99-1474-359 of  
 SEQ ID NOs 240 and 317; 99-1479-158 of SEQ ID NOs 241 and 318; 99-1479-379 of SEQ ID  
 NOs 242 and 319; 99-148-129 of SEQ ID NOs 243 and 320; 99-148-132 of SEQ ID NOs 244  
 and 321; 99-148-139 of SEQ ID NOs 245 and 322;  
 20 99-148-140 of SEQ ID NOs 246 and 323; 99-148-182 of SEQ ID NOs 247 and 324;  
 99-148-366 of SEQ ID NOs 248 and 325; 99-148-76 of SEQ ID NOs 249 and 326;  
 99-1480-290 of SEQ ID NOs 250 and 327; 99-1481-285 of SEQ ID NOs 251 and 328; 99-  
 1484-101 of SEQ ID NOs 252 and 329; 99-1484-328 of SEQ ID NOs 253 and 330; 99-1485-  
 251 of SEQ ID NOs 254 and 331; 99-1490-381 of SEQ ID NOs 255 and 332; 99-1493-280 of  
 25 SEQ ID NOs 256 and 333; 99-151-94 of SEQ ID NOs 257 and 334;  
 99-211-291 of SEQ ID NOs 258 and 335; 99-213-37 of SEQ ID NOs 259 and 336;  
 99-221-442 of SEQ ID NOs 260 and 337; 99-222-109 of SEQ ID NOs 261 and 338; and the  
 compliments thereof.

30 The term “non-genic” is used herein to describe PG1-related biallelic markers, as well  
 as polynucleotides and primers which do not occur in the human PG1 genomic sequence of  
 SEQ ID NO: 179. The term “genic” is used herein to describe PG1-related biallelic markers as  
 well as polynucleotides and primers which do occur in the human PG1 genomic sequence of  
 SEQ ID NO: 179.

Kawana Y, Killary AM, Rinker-Schaeffer CW, Barrett JC, Isaacs JT, Kugoh H, Oshimura M, Shimazaki J, Prostate Suppl. 1996; 6: 31-35).

Recently Washburn et al. were able to find substantial numbers of tumors with the allelic loss specific to 8p23 by LOH studies of 31 cases of human prostate cancer. (Washburn J, Woino K, and Macoska J, Proceedings of American Association for Cancer Research, March 1997; 38). In these samples they were able to define the minimal overlapping region with deletions covering genetic interval D8S262-D8S277.

Linkage Analysis Studies: Search for Prostate Cancer

Linked Regions on Chromosome 8

Microsatellite markers mapping to chromosome 8 were used by the inventors to perform linkage analysis studies on 194 individuals issued from 47 families affected with prostate cancer. While multiple point analysis led to weak linkage results, two point lod score analysis led to non significant results, as shown below.

*Two point lod (parametric analysis)*

<i>MARKER</i>	<i>Distance (cM)</i>	<i>Z(lod) scores</i>
<i>D8S1742</i>	<i>0.8</i>	<i>-0.13</i>
<i>D8S561</i>		<i>-0.07</i>

# of families analyzed	47
Total # of individuals genotyped	194
Total # of affected individuals genotyped	122

Example 3 describes the preparation of genomic DNA samples from the individuals screened to identify biallelic markers.

### Example 3

The population used in order to generate biallelic markers in the region of interest consisted of ca. 100 unrelated individuals corresponding to a French heterogeneous population.

DNA was extracted from peripheral venous blood of each donor as follows.

30 ml of blood were taken in the presence of EDTA. Cells (pellet) were collected after centrifugation for 10 minutes at 2000 rpm. Red cells were lysed by a lysis solution (50 ml final volume : 10 mM Tris pH7.6; 5 mM MgCl<sub>2</sub>; 10 mM NaCl). The solution was centrifuged (10 minutes, 2000 rpm) as many times as necessary to eliminate the residual red cells present in the supernatant, after resuspension of the pellet in the lysis solution.

The pellet of white cells was lysed overnight at 42°C with 3.7 ml of lysis solution composed of:

- 3 ml TE 10-2 (Tris-HCl 10 mM, EDTA 2 mM) / NaCl 0.4 M
- 200 µl SDS 10%
- 500 µl K-proteinase (2 mg K-proteinase in TE 10-2 / NaCl 0.4 M).

For the extraction of proteins, 1 ml saturated NaCl (6M) (1/3.5 v/v) was added. After vigorous agitation, the solution was centrifuged for 20 minutes at 10000 rpm.

For the precipitation of DNA, 2 to 3 volumes of 100% ethanol were added to the previous supernatant, and the solution was centrifuged for 30 minutes at 2000 rpm. The DNA solution was rinsed three times with 70% ethanol to eliminate salts, and centrifuged for 20 minutes at 2000 rpm. The pellet was dried at 37°C, and resuspended in 1 ml TE 10-1 or 1 ml water. The DNA concentration was evaluated by measuring the OD at 260 nm (1 unit OD = 50 µg/ml DNA).

To determine the presence of proteins in the DNA solution, the OD 260 / OD 280 ratio was determined. Only DNA preparations having a OD 260 / OD 280 ratio between 1.8 and 2 were used in the subsequent steps described below.

### DNA Amplification

Once each BAC was isolated, pairs of primers, each one defining a 500 bp-amplification fragment, were designed. Each of the primers contained a common oligonucleotide tail upstream of the specific bases targeted for amplification, allowing the amplification products from each set of primers to be sequenced using the common sequence as a sequencing primer. The primers used for the genomic amplification of sequences derived from BACs were defined with the QSP software (Hillier L. and Green P. Methods Appl., 1991,

1: 124-8). The synthesis of primers was performed following the phosphoramidite method, on a GENSET UFPS 24.1 synthesizer.

Example 4 provides the procedures used in the amplification reactions.

#### Example 4

The amplification of each sequence was performed by PCR (Polymerase Chain Reaction) as follows:

- final volume	50 $\mu$ l
- genomic DNA	100 ng
- MgCl <sub>2</sub>	2 mM
- dNTP (each)	200 $\mu$ M
- primer (each)	7.5 pmoles
- Ampli Taq Gold DNA polymerase (Perkin)	1 unit
- PCR buffer (10X=0.1 M Tris HCl pH 8.3, 0.5 M KCl) 1X.	

The amplification was performed on a Perkin Elmer 9600 Thermocycler or MJ Research PTC200 with heating lid. After heating at 94°C for 10 minutes, 35 cycles were performed. Each cycle comprised: 30 sec at 94°C, 1 minute at 55°C, and 30 sec at 72°C. For final elongation, 7 minutes at 72°C ended the amplification.

The obtained quantity of amplification products was determined on 96-well microtiter plates, using a fluorimeter and Picogreen as intercalating agent (Molecular Probes).

The sequences of the amplification products were determined for each of the approximately 100 individuals from whom genomic DNA was obtained. Those amplification products which contained biallelic markers were identified.

Figure 1 shows the locations of the biallelic markers along the 8p23 BAC contig. This first set of markers corresponds to a medium density map of the candidate locus, with an inter-marker distance averaging 50kb-150kb.

A second set of biallelic markers was then generated as described above in order to provide a very high-density map of the region identified using the first set of markers which can be used to conduct association studies, as explained below. The high density map has markers spaced on average every 2-50kb.

The biallelic markers were then used in association studies as described below.

#### Collection of DNA samples from affected and non-affected individuals

Prostate cancer patients were recruited according to clinical inclusion criteria based on pathological or radical prostatectomy records. Control cases included in this study were both ethnically- and age-matched to the affected cases; they were checked for both the absence of all

(Amersham), both of the fluorescent ddNTPs (Perkin Elmer) corresponding to the polymorphism {0.025  $\mu$ l ddTTP and ddCTP, 0.05  $\mu$ l ddATP and ddGTP}, H<sub>2</sub>O to a final volume of 20  $\mu$ l. A PCR program on a GeneAmp 9600 thermocycler was carried out as follows: 4 minutes at 94°C ; 5 sec at 55°C / 10 sec at 94°C for 20 cycles. The reaction product was incubated at 4°C until precipitation. The microtiter plate was centrifuged 10 sec at 1500 rpm. 19  $\mu$ l MgCl<sub>2</sub> 2mM and 55  $\mu$ l 100 % ethanol were added in each well. After 15 minute incubation at room temperature, the microtiter plate was centrifuged at 3300 rpm 15 minutes at 4°C. Supernatants were discarded by inverting the microtitre plate on a box folded to proper size and by centrifugation at 300 rpm 2 minutes at 4°C afterwards. The microplate was then dried 5 minutes in a vacuum drier. The pellets were resuspended in 2.5  $\mu$ l formamide EDTA loading buffer (0.7 $\mu$ l of 9  $\mu$ g/ $\mu$ l dextran blue in 25 mM EDTA and 1.8  $\mu$ l formamide). A 10% polyacrylamide gel / 12 cm / 64 wells was pre-run for 5 minutes on a 377 ABI 377 sequencer. After 5 minutes denaturation at 100°C, 0.8  $\mu$ l of each microsequencing reaction product was loaded in each well of the gel. After migration (2 h 30 for 2 microtiter plates of PCR products per gel), the fluorescent signals emitted by the incorporated ddNTPs were analyzed on the ABI 377 sequencer using the GENESCAN software (Perkin Elmer). Following gel analysis, data were automatically processed with a software that allowed the determination of the alleles of biallelic markers present in each amplified fragment.

#### **I.D. Initial Association Studies**

Association studies were run in two successive steps. In a first step, a rough localization of the candidate gene was achieved by determining the frequencies of the biallelic markers of Figure 1 in the affected and unaffected populations. The results of this rough localization are shown in Figure 2. This analysis indicated that a gene responsible for prostate cancer was located near the biallelic marker designated 4-67.

In a second phase of the analysis, the position of the gene responsible for prostate cancer was further refined using the very high density set of markers described above. The results of this localization are shown in Figure 3.

As shown in Figure 3, the second phase of the analysis confirmed that the gene responsible for prostate cancer was near the biallelic marker designated 4-67, most probably within a ca. 150kb region comprising the marker.

#### **Haplotype analysis**

The allelic frequencies of each of the alleles of biallelic markers 99-123, 4-26, 4-14, 4-77, 99-217, 4-67, 99-213, 99-221, and 99-135 (SEQ ID NOs: 21-38) were determined in the affected and unaffected populations. Table 1 lists the internal identification numbers of the

indicates that the transformation process necessary to the generation of these normal prostatic cell lines might result in similar alteration which further support the previous hypothesis.

#### Example 9

#### Determining the Tumor Suppressor Activity of the PG1 Gene Product, Mutants and Other PG1

5

#### Polypeptides

PG1 variants which results from either alternate splicing of the PG1 mRNA or from mutation of PG1 that introduce a stop codon (nucleotide of SEQ ID NO: 69 and protein of SEQ ID NO: 70) can no longer perform its role of tumor suppressor. It is possible and even likely that PG1 tumor suppressor role extends beyond prostate cancer to other form of malignancies. PG1 therefore represent a prime candidate for gene therapy of cancer by creating a targeting vector which knocks out the mutant and/or introduces a wild-type PG1 gene (e.g. SEQ ID NO 3 or 179) or a fragment thereof.

To validate this model, PG1 and its alternatively spliced or mutated variants are stably transfected in tumor cell line using methods described in Section VIII. The efficiency of transfection is determined by northern and western blotting; the latter is performed using antibodies prepared against PG1 synthetic peptides designed to distinguish the product of the most abundant PG1 mRNA from the alternatively spliced variants, the truncated variant, or other functional mutants. The production of synthetic peptides and of polyclonal antibodies is performed using the methods described herein in Sections III and VII. After demonstrating that PG1 and its variant are efficiently expressed in various tumor cell line preferably derived from human prostate cancer, hepatocarcinoma, lung and colon carcinoma; we the effect of this gene on the rate of cell division, DNA synthesis, ability to grow in soft agar and ability to induce tumor progression and metastasis when injected in immunologically deficient nude mice are determined.

Alternatively the PG1 gene and its variant are inserted in adenoviruses that are used to obtain a high level of expression of these genes. This method is preferred to test the effect of PG1 expression in animal that are spontaneously developing tumor. The production of specific adenoviruses is obtained using methods familiar to those with normal skills in cell and molecular biology.

30

#### II. POLYNUCLEOTIDES:

The present invention encompasses polynucleotides in the form of PG1 genomic or cDNA as well as polynucleotides for use as primers and probes in the methods of the invention. These polynucleotides may consist of, consist essentially of, or comprise a contiguous span of

Primers with their 3' ends located 1 nucleotide upstream or downstream of a PG1-related biallelic marker have a special utility in microsequencing assays. Preferred microsequencing primers include the polynucleotides from position 1 to position 23 and from position 25 to position 47 of SEQ ID NOs: 21-38, and as well as the compliments thereof.

5 Additional preferred microsequencing primers for particular non-genic PG1-related biallelic markers are listed as follows by the internal reference number for the marker and the SEQ ID NOs of the two preferred microsequencing primers:

4-14-107 of SEQ ID NOs 425 and 502\*; 4-14-317 of SEQ ID NOs 426 and 503\*;  
4-14-35 of SEQ ID NOs 427 and 504\*; 4-20-149 of SEQ ID NOs 428\* and 505;  
10 4-20-77 of SEQ ID NOs 429 and 506; 4-22-174 of SEQ ID NOs 430\* and 507;  
4-22-176 of SEQ ID NOs 431 and 508; 4-26-60 of SEQ ID NOs 432 and 509\*;  
4-26-72 of SEQ ID NOs 433 and 510; 4-3-130 of SEQ ID NOs 434 and 511\*;  
4-38-63 of SEQ ID NOs 435 and 512; 4-38-83 of SEQ ID NOs 436 and 513\*;  
4-4-152 of SEQ ID NOs 437 and 514; 4-4-187 of SEQ ID NOs 438\* and 515;  
15 4-4-288 of SEQ ID NOs 439 and 516; 4-42-304 of SEQ ID NOs 440 and 517;  
4-42-401 of SEQ ID NOs 441\* and 518; 4-43-328 of SEQ ID NOs 442 and 519;  
4-43-70 of SEQ ID NOs 443\* and 520; 4-50-209 of SEQ ID NOs 444\* and 521;  
4-50-293 of SEQ ID NOs 445\* and 522; 4-50-323 of SEQ ID NOs 446\* and 523;  
4-50-329 of SEQ ID NOs 447\* and 524; 4-50-330 of SEQ ID NOs 448 and 525;  
20 4-52-163 of SEQ ID NOs 449\* and 526; 4-52-88 of SEQ ID NOs 450\* and 527;  
4-53-258 of SEQ ID NOs 451 and 528\*; 4-54-283 of SEQ ID NOs 452\* and 529;  
4-54-388 of SEQ ID NOs 453 and 530; 4-55-70 of SEQ ID NOs 454 and 531\*;  
4-55-95 of SEQ ID NOs 455\* and 532; 4-56-159 of SEQ ID NOs 456\* and 533;  
4-56-213 of SEQ ID NOs 457 and 534; 4-58-289 of SEQ ID NOs 458\* and 535;  
25 4-58-318 of SEQ ID NOs 459\* and 536; 4-60-266 of SEQ ID NOs 460\* and 537;  
4-60-293 of SEQ ID NOs 461\* and 538; 4-84-241 of SEQ ID NOs 462 and 539\*;  
4-84-262 of SEQ ID NOs 463 and 540; 4-86-206 of SEQ ID NOs 464 and 541\*;  
4-86-309 of SEQ ID NOs 465 and 542; 4-88-349 of SEQ ID NOs 466 and 543.;  
4-89-87 of SEQ ID NOs 467\* and 544.; 99-123-184 of SEQ ID NOs 468 and 545;  
30 99-128-202 of SEQ ID NOs 469 and 546; 99-128-275 of SEQ ID NOs 470 and 547;  
99-128-313 of SEQ ID NOs 471 and 548; 99-128-60 of SEQ ID NOs 472\* and 549;  
99-12907-295 of SEQ ID NOs 473 and 550\*;  
99-130-58 of SEQ ID NOs 474\* and 551\*;  
99-134-362 of SEQ ID NOs 475 and 552\*; 99-140-130 of SEQ ID NOs 476\* and 553\*;

- 99-1462-238 of SEQ ID NOs 477\* and 554; 99-147-181 of SEQ ID NOs 478 and 555\*;  
99-1474-156 of SEQ ID NOs 479 and 556\*; 99-1474-359 of SEQ ID NOs 480 and 557;  
99-1479-158 of SEQ ID NOs 481\* and 558; 99-1479-379 of SEQ ID NOs 482 and 559;  
99-148-129 of SEQ ID NOs 483 and 560; 99-148-132 of SEQ ID NOs 484 and 561;  
5 99-148-139 of SEQ ID NOs 485 and 562; 99-148-140 of SEQ ID NOs 486 and 563;  
99-148-182 of SEQ ID NOs 487 and 564\*; 99-148-366 of SEQ ID NOs 488 and 565;  
99-148-76 of SEQ ID NOs 489 and 566; 99-1480-290 of SEQ ID NOs 490 and 567\*;  
99-1481-285 of SEQ ID NOs 491 and 568\*; 99-1484-101 of SEQ ID NOs 492 and 569;  
99-1484-328 of SEQ ID NOs 493\* and 570;  
10 99-1485-251 of SEQ ID NOs 494 and 571\*;  
99-1490-381 of SEQ ID NOs 495\* and 572;  
99-1493-280 of SEQ ID NOs 496 and 573\*;  
99-151-94 of SEQ ID NOs 497 and 574\*; 99-211-291 of SEQ ID NOs 498\* and 575;  
99-213-37 of SEQ ID NOs 499 and 576; 99-221-442 of SEQ ID NOs 500 and 577;  
15 99-222-109 of SEQ ID NOs 501\* and 578; and compliments thereof.

Additional preferred microsequencing primers for particular genic PG1-related biallelic markers include a polynucleotide selected from the group consisting of the nucleotide sequences from position N-X to position N-1 of SEQ ID NO:179, nucleotide sequences from position N+1 to position N+X of SEQ ID NO:179, and the compliments thereof, wherein X is  
20 equal to 15, 18, 20, 25, 30, or a range of 15 to 30, and N is equal to one of the following values:  
2159; 2443; 4452; 5733; 8438; 11843; 1983; 12080; 12221; 12947; 13147; 13194; 13310;  
13342; 13367; 13594; 13680; 13902; 16231; 16388; 17608; 18034; 18290; 18786; 22835;  
22872; 25183; 25192; 25614; 26911; 32703; 34491; 34756; 34934; 5160; 39897; 40598;  
40816; 40947; 45783; 47929; 48206; 48207; 49282; 50037; 50054; 50101; 50220; 50440;  
25 50562; 50653; 50660; 50745; 50885; 51249; 51333; 51435; 51468; 51515; 51557; 51566;  
51632; 51666; 52016; 52096; 52151; 52282; 52348; 52410; 52580; 52712; 52772; 52860;  
53092; 53272; 53389; 53511; 53600; 53665; 53815; 54365; and 54541.

The probes of the present invention is designed from the disclosed sequences for any method known in the art, particularly methods which allow for testing if a particular sequence  
30 or marker disclosed herein is present. A preferred set of probes is designed for use in the hybridization assays of the invention in any manner known in the art such that they selectively bind to one allele of a biallelic marker, but not the other under any particular set of assay conditions. Preferred hybridization probes may consists of, consist essentially of, or comprise a contiguous span which ranges in length from 8, 10, 12, 15, 18 or 20 to 25, 35, 40, 50, 60, 70, or

the Cre enzyme directly into the desired cell, such as described by Araki K. et al., 1995, Proc. Natl. Acad. Sci. USA, 92: 160-164; or by lipofection of the enzyme into the cells, such as described by Baubonis et al., 1993, Nucleic Acids Res., 21: 2025-2029; (b) transfecting the cell host with a vector comprising the Cre coding sequence operably linked to a promoter functional in the recombinant cell host, which promoter being optionally inducible, said vector being introduced in the recombinant cell host, such as described by Gu H. et al., 1993, Cell, 73: 1155-1164; and Sauer B. et al., 1988, Proc. Natl. Acad. Sci. USA, 85: 5166-5170; (c) introducing in the genome of the host cell a polynucleotide comprising the Cre coding sequence operably linked to a promoter functional in the recombinant cell host, which promoter is optionally inducible, and said polynucleotide being inserted in the genome of the cell host either by a random insertion event or an homologous recombination event, such as described by Gu H. et al., 1994, Science, 265: 103-106.

In the specific embodiment wherein the vector containing the sequence to be inserted in the PG1 gene by homologous recombination is constructed in such a way that selectable markers are flanked by loxP sites of the same orientation, it is possible, by treatment by the Cre enzyme, to eliminate the selectable markers while leaving the PG1 sequences of interest that have been inserted by an homologous recombination event. Again, two selectable markers are needed: a positive selection marker to select for the recombination event and a negative selection marker to select for the homologous recombination event. Vectors and methods using the Cre-loxP system are described by Zou Y.R. et al., 1994, Curr. Biol., 4: 1099-1103.

Thus, a third preferred DNA construct of the invention comprises, from 5'-end to 3'-end: (a) a first nucleotide sequence that is comprised of a PG1 sequence, preferably a PG1 genomic sequence; (b) a nucleotide sequence comprising a polynucleotide encoding a positive selection marker, such as the marker for neomycin resistance (neo), said nucleotide sequence comprising additionally two sequences defining a site recognized by a recombinase, such as a loxP site, the two sites being placed in the same orientation; and (c) a second nucleotide sequence that is comprised of a PG1 sequence, preferably a PG1 genomic sequence, and is located on the genome downstream of the first PG1 nucleotide sequence (a).

The sequences defining a site recognized by a recombinase, such as a loxP site, are preferably located within the nucleotide sequence (b) at suitable locations bordering the nucleotide sequence for which the conditional excision is sought. In one specific embodiment, two loxP sites are located at each side of the positive selection marker sequence, in order to allow its excision at a desired time after the occurrence of the homologous recombination event.

### Expression Vectors

The polynucleotides of the invention also include expression vectors. Expression vector systems, control sequences and compatible host are known in the art. For a review of these systems see, for example, U.S. Patent No. 5,350,671, columns 45-48. Any of the standard methods known to those skilled in the art for the insertion of DNA fragments into a vector is used to construct expression vectors containing a chimeric gene consisting of appropriate transcriptional/translational control signals and the protein coding sequences. These methods may include in vitro recombinant DNA and synthetic techniques and in vivo recombinants (genetic recombination).

Expression of a polypeptide, peptide or derivative, or analogs thereof encoded by a polynucleotide sequence in SEQ ID NOs: 3, 69, 100-112, or 179-184 is regulated by a second nucleic acid sequence so that the protein or peptide is expressed in a host transformed with the recombinant DNA molecule. For example, expression of a protein or peptide is controlled by any promoter/enhancer element known in the art. Promoters which is used to control expression include, but are not limited to, the CMV promoter, the SV40 early promoter region (Bernoist and Chambon, 1981, *Nature* 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto, et al., 1980, *Cell* 22:787-797), the herpes thymidine kinase promoter (Wagner et al., 1981, *Proc. Natl. Acad. Sci. U.S.A.* 78:1441-1445), the regulatory sequences of the metallothionein gene (Brinster et al., 1982, *Nature* 296:39-42); prokaryotic expression vectors such as the beta-lactamase promoter (Villa-Kamaroff, et al., 1978, *Proc. Natl. Acad. Sci. U.S.A.* 75:3727-3731), or the tac promoter (DeBoer, et al., 1983, *Proc. Natl. Acad. Sci. U.S.A.* 80:21-25); see also "Useful proteins from recombinant bacteria" in *Scientific American*, 1980, 242:74-94; plant expression vectors comprising the nopaline synthetase promoter region (Herrera-Estrella et al., 1983, *Nature* 303:209-213) or the cauliflower mosaic virus 35S RNA promoter (Gardner, et al., 1981, *Nucl. Acids Res.* 9:2871), and the promoter of the photosynthetic enzyme ribulose biphosphate carboxylase (Herrera-Estrella et al., 1984, *Nature* 310:115-120); promoter elements from yeast or other fungi such as the Gal 4 promoter, the ADC (alcohol dehydrogenase) promoter, PGK (phosphoglycerol kinase) promoter, alkaline phosphatase promoter, and the following animal transcriptional control regions, which exhibit tissue specificity and have been utilized in transgenic animals: elastase I gene control region which is active in pancreatic acinar cells (Swift et al., 1984, *Cell* 38:639-646; Ornitz et al., 1986, *Cold Spring Harbor Symp. Quant. Biol.* 50:399-409; MacDonald, 1987, *Hepatology* 7:425-515); insulin gene control region which is active in pancreatic beta cells (Hanahan, 1985, *Nature* 315:115-122), immunoglobulin gene control

region which is active in lymphoid cells (Grosschedl et al., 1984, Cell 38:647-658; Adames et al., 1985, Nature 318:533-538; Alexander et al., 1987, Mol. Cell. Biol. 7:1436-1444), mouse mammary tumor virus control region which is active in testicular, breast, lymphoid and mast cells (Leder et al., 1986, Cell 45:485-495), albumin gene control region which is active in liver (Pinkert et al., 1987, Genes and Devel. 1:268-276), alpha-fetoprotein gene control region which is active in liver (Krumlauf et al., 1985, Mol. Cell. Biol. 5:1639-1648; Hammer et al., 1987, Science 235:53-58; alpha 1-antitrypsin gene control region which is active in the liver (Kelsey et al., 1987, Genes and Devel. 1:161-171), beta-globin gene control region which is active in myeloid cells (Mogram et al., 1985, Nature 315:338-340; Kollias et al., 1986, Cell 46:89-94; myelin basic protein gene control region which is active in oligodendrocyte cells in the brain (Readhead et al., 1987, Cell 48:703-712); myosin light chain-2 gene control region which is active in skeletal muscle (Sani, 1985, Nature 314:283-286), and gonadotropic releasing hormone gene control region which is active in the hypothalamus (Mason et al., 1986, Science 234:1372-1378).

Other suitable vectors, particularly for the expression of genes in mammalian cells, is selected from the group of vectors consisting of P1 bacteriophages, and bacterial artificial chromosomes (BACs). These types of vectors may contain large inserts ranging from about 80-90 kb (P1 bacteriophage) to about 300 kb (BACs).

#### P1 bacteriophage

The construction of P1 bacteriophage vectors such as p158 or p158/neo8 are notably described by Sternberg N.L., 1992, Trends Genet., 8: 1-16; and Sternberg N.L., 1994, Mamm. Genome, 5: 397-404. Recombinant P1 clones comprising PG1 nucleotide sequences is designed for inserting large polynucleotides of more than 40 kb (Linton M.F. et al., 1993, J. Clin. Invest., 92: 3029-3037). To generate P1 DNA for transgenic experiments, a preferred protocol is the protocol described by McCormick et al., 1994, Genet. Anal. Tech. Appl., 11: 158-164. Briefly, E. coli (preferably strain NS3529) harboring the P1 plasmid are grown overnight in a suitable broth medium containing 25 µg/ml of kanamycin. The P1 DNA is prepared from the E. coli by alkaline lysis using the Qiagen Plasmid Maxi kit (Qiagen, Chatsworth, CA, USA), according to the manufacturer's instructions. The P1 DNA is purified from the bacterial lysate on two Qiagen-tip 500 columns, using the washing and elution buffers contained in the kit. A phenol/chloroform extraction is then performed before precipitating the DNA with 70% ethanol. After solubilizing the DNA in TE (10 mM Tris-HCl, pH 7.4, 1 mM EDTA), the concentration of the DNA is assessed by spectrophotometry.

To study the function of PG1, it is inactivate by homologous recombination in the remaining allele of PG1 in the PC3 cell line. To inactivate the remaining PG1 allele, a knock-out targeting vector is generated by inserting two genomic DNA fragments of 3.0 and 4.3 kb (that correspond to a sequence upstream of the PG1 promoter and to part of intron 1, respectively) in the pKO Scrambler Neo TK vector (Lexicon ref V1901). Since the targeting vector contains the neomycine resistance gene as well as the Tk gene, homologous recombination is selected by adding geneticin and FIAU to the medium. The promoter, the transcriptional start site, and the first ATG contained in exon 1 on the recombinant allele is deleted by homologous recombination between the targeting vector and the remaining PG1 allele. Accordingly, no coding transcripts is initiated from the recombinant allele. The parental PC3 cells as well as cells hemizygous for the null allele are assessed for their phenotype, their growth rate in liquid culture, their ability to grow in agar (anchorage-independent growth) as well as their ability to form tumors and metastasis when injected subcutaneously in nude mice.

To determine the function of PG1 in the animal, and to generate an animal model for prostate tumorigenesis, mice in which tissue specific inactivation of the PG1 alleles can be induced are generated. For this purpose, the Cre-loxP system is utilized as described above to allow chromosome engineering to be perform directly in the animal.

First, to generate mice with a conditional null allele, two loxP sites are introduced in the murine genome, the first one 5' to the PG1 promoter and the second one 3' to the PG1 exon 1. Alternatively, to generate subtle mutations or to specifically mutate some isoforms, the loxP sites are introduced so that they flank any of the given exons or any potential set of exons. It is important to note that a functional PG1 messenger can be transcribed from these alleles until a recombination is triggered between the loxP sites by the Cre enzyme.

Second, to generate the inducer mice, the Cre gene is introduced in the mouse genome under the control of a tissue specific promoter, for example under the control of the PSA (prostate specific antigen) promoter.

Finally, tissue specific inactivation of the PG1 gene are induced by generating mice containing the Cre transgene that are homozygous for the recombinant PG1 allele.

#### Gene Therapy

The present invention also comprises the use of the PG1 genomic DNA sequence of SEQ ID NO: 179, the PG1 cDNA of SEQ ID NO: 3, or nucleic acid encoding a mutant PG1 protein responsible for a detectable phenotype in gene therapy strategies, including antisense and triple helix strategies as described in Examples 19 and 20, below. In antisense approaches, nucleic acid sequences complementary to an mRNA are hybridized to the mRNA intracellularly, thereby

alternatively be present on the template. For example, immobilization can be carried out via an interaction between biotinylated DNA and streptavidin-coated microtitration wells or avidin-coated polystyrene particles. In the same manner oligonucleotides or templates is attached to a solid support in a high-density format. In such solid phase microsequencing reactions, incorporated ddNTPs can be radiolabeled (Syvänen, *Clinica Chimica Acta* 226:225-236, 1994) or linked to fluorescein (Livak and Hainer, *Human Mutation* 3:379-385, 1994). The detection of radiolabeled ddNTPs can be achieved through scintillation-based techniques. The detection of fluorescein-linked ddNTPs can be based on the binding of anti-fluorescein antibody conjugated with alkaline phosphatase, followed by incubation with a chromogenic substrate (such as *p*-nitrophenyl phosphate). Other possible reporter-detection pairs include: ddNTP linked to dinitrophenyl (DNP) and anti-DNP alkaline phosphatase conjugate (Harju et al., *Clin. Chem.* 39/11 2282-2287, 1993) or biotinylated ddNTP and horseradish peroxidase-conjugated streptavidin with *o*-phenylenediamine as a substrate (WO 92/15712). As yet another alternative solid-phase microsequencing procedure, Nyren et al. (*Analytical Biochemistry* 208:171-175, 1993) described a method relying on the detection of DNA polymerase activity by an enzymatic luminometric inorganic pyrophosphate detection assay (ELIDA).

Pastinen et al. (*Genome research* 7:606-614, 1997) describe a method for multiplex detection of single nucleotide polymorphism in which the solid phase minisequencing principle is applied to an oligonucleotide array format. High-density arrays of DNA probes attached to a solid support (DNA chips) are further described in X.C.5.

In one aspect the present invention provides polynucleotides and methods to genotype one or more biallelic markers of the present invention by performing a microsequencing assay. It will be appreciated that any primer having a 3' end immediately adjacent to the polymorphic nucleotide is used. However, polynucleotides comprising at least 8, 12, 15, 20, 25, or 30 consecutive nucleotides of the sequence immediately adjacent to the biallelic marker and having a 3' terminus immediately upstream of the corresponding biallelic marker are well suited for determining the identity of a nucleotide at biallelic marker site.

Similarly, it will be appreciated that microsequencing analysis is performed for any biallelic marker or any combination of biallelic markers of the present invention.

### **Mismatch detection assays based on polymerases and ligases**

In one aspect the present invention provides polynucleotides and methods to determine the allele of one or more biallelic markers of the present invention in a biological sample, by mismatch detection assays based on polymerases and/or ligases. These assays are based on the specificity of polymerases and ligases. Polymerization reactions places particularly stringent

150, 200, 500, or 1000 nucleotides of SEQ ID No 3 or the complements thereof, wherein said contiguous span comprises at least 1, 2, 3, 5, 10, or 25 of the following nucleotide positions of SEQ ID No 3: 1-280, 651-690, 3315-4288, and 5176-5227; and c) a nucleotide sequence complementary to either one of the preceding nucleotide sequences.

5           The "nucleic acid codes of the invention" further encompass nucleotide sequences homologous to: a) a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 179, wherein said contiguous span comprises at least 1, 2, 3, 5, 10, or 25 of the following nucleotide positions of SEQ ID No 179: 1-2324, 2852-2936, 3204-3249, 3456-3572, 3899-4996, 5028-6086, 6310-8710, 9136-11170, 10   11534-12104, 12733-13163, 13206-14150, 14191-14302, 14338-14359, 14788-15589, 16050-16409, 16440-21718, 21959-22007, 22086-23057, 23488-23712, 23832-24099, 24165-24376, 24429-24568, 24607-25096, 25127-25269, 25300-27576, 27612-29217, 29415-30776, 30807-30986, 31628-32658, 32699-36324, 36772-39149, 39184-40269, 40580-40683, 40844-41048, 41271-43539, 43570-47024, 47510-48065, 48192-49692, 49723-50174, 52626-53599, 54516-15   55209, and 55666-56146; b) a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 3 or the complements thereof, wherein said contiguous span comprises at least 1, 2, 3, 5, 10, or 25 of the following nucleotide positions of SEQ ID No 3: 1-280, 651-690, 3315-4288, and 5176-5227; and, c) sequences complementary to all of the preceding sequences. Homologous sequences refer to a sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, 80%, or 75% homology to these contiguous spans. Homology may be determined using any method described herein, including BLAST2N with the default parameters or with any modified parameters. Homologous sequences also may include RNA sequences in which uridines replace the thymines in the nucleic acid codes of the invention. It will be appreciated that the nucleic acid codes of the invention can be represented in the traditional single character format (See the inside back cover of Stryer, Lubert. Biochemistry, 3<sup>rd</sup> edition. W. H Freeman & Co., New York.) or in any other format or code which records the identity of the nucleotides in a sequence.

25           As used herein the term "polypeptide codes of the invention" encompass the polypeptide sequences comprising a contiguous span of at least 6, 8, 10, 12, 15, 20, 25, 30, 40, 30   50, or 100 amino acids of SEQ ID No 4, wherein said contiguous span includes at least 1, 2, 3, or 5 of the amino acid positions 1-26, 295-302, and 333-353. It will be appreciated that the polypeptide codes of the invention can be represented in the traditional single character format or three letter format (See the inside back cover of Stryer, Lubert. Biochemistry, 3<sup>rd</sup> edition. W. H

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 7,189,833

Page 1 of 2

APPLICATION NO.: 09/901,484

DATED : March 13, 2007

INVENTORS : Daniel Cohen, Marta Blumenfeld, Ilya Chumakov, Lydie Bougueleret

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 19,

Line 12, "The tern" should read --The term--.

Line 53, "and 35" should read --and 315--.

Column 22,

Line 22, "Washburn et al." should read --Washburn et al.--.

Lines 24-25, "(Washburn)" should read --(Washburn--.

Column 29,

Line 28, "5 MM MgCl<sub>2</sub>" should read --5 mM MgCl<sub>2</sub>--.

Column 30,

Line 25, "72° C. final elongation" should read --72° C. For final elongation--.

Column 31,

Line 45, "The re was incubated" should read --The reaction product was incubated--.

Column 47,

Line 14, "Tumor Supressor" should read --Tumor Suppressor--.

MAILING ADDRESS OF SENDER:

Saliwanchik, Lloyd & Saliwanchik  
P.O. Box 142950  
Gainesville, FL 32614-2950

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 7,189,833

Page 2 of 2

APPLICATION NO.: 09/901,484

DATED : March 13, 2007

INVENTORS : Daniel Cohen, Marta Blumenfeld, Ilya Chumakov, Lydie Bougueleret

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 51.

Line 10, "44-187" should read -- 4-4-187--.

Line 12, "442-304" should read -- 4-42-304--.

Line 13, "443-328" should read -- 4-43-328--.

Line 28, "SEQ ID NOs 464" should read --SEQ ID NOs 464--.

Line 54, "99-221442 of" should read --99-221-442 of--.

Column 82.

Line 37, "by loxp sites" should read --by loxP sites--.

Column 84.

Line 6, "(Bemoist" should read --(Bernoist--.

Column 85.

Line 4, "(Stemberg N. L., 1994" should read --(Sternberg N.L., 1994--.

Column 89.

Line 56, "the loxp sites" should read --the loxP sites--.

Column 101.

Lines 59-60, "with 0-phenylenediamine" should read --with *o*-phenylenediamine--.

Column 123.

Line 9, "nucteic acid" should read --nucleic acid--.

MAILING ADDRESS OF SENDER:

Saliwanchik, Lloyd & Saliwanchik

P.O. Box 142950

Gainesville, FL 32614-2950